

GenCore version 4.5
copyright: (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:18 ; Search time 210.42 Seconds

(without alignments)

BLOSUM62 7.638 Million cell updates/sec

title: US-09-331-631A-7_COPY_34_80
Perfect score: 258
Sequence: YERDPRQQYEQCORCESEA.....QCEQREREYKEQQRQQEE 47

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36:*

1: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1980.DAT: *
2: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1981.DAT: *
3: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1982.DAT: *
4: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1983.DAT: *
5: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1985.DAT: *
6: /SIDS1/gcdata/geneseq/geneseq/AA1986.DAT: *
7: /SIDS1/gcdata/geneseq/geneseq/AA1987.DAT: *
9: /SIDS1/gcdata/geneseq/geneseq/AA1989.DAT: *
10: /SIDS1/gcdata/geneseq/geneseq/AA1990.DAT: *
11: /SIDS1/gcdata/geneseq/geneseq/AA1991.DAT: *
12: /SIDS1/gcdata/geneseq/geneseq/AA1992.DAT: *
13: /SIDS1/gcdata/geneseq/geneseq/AA1993.DAT: *
14: /SIDS1/gcdata/geneseq/geneseq/AA1994.DAT: *
15: /SIDS1/gcdata/geneseq/geneseq/AA1995.DAT: *
17: /SIDS1/gcdata/geneseq/geneseq/AA1996.DAT: *
18: /SIDS1/gcdata/geneseq/geneseq/AA1997.DAT: *
19: /SIDS1/gcdata/geneseq/geneseq/AA1998.DAT: *
20: /SIDS1/gcdata/geneseq/geneseq/AA1999.DAT: *
21: /SIDS1/gcdata/geneseq/geneseq/AA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT	ID	Description
1	W62831	standard; Protein; 525 AA.
	W62831	AC
	XX	DT
	XX	27-OCT-1998 (first entry)
	XX	Theobroma cacao antimicrobial protein.
	XX	KW antimicrobial protein; infestation; control.
	XX	OS Theobroma cacao.
	XX	PN W09877805-A1.
	XX	PD 02-JUL-1998.
	XX	PP
	XX	22-DEC-1997; 97WO-AU000874.
	PR	20-DEC-1996; 96AU-0004275.
	XX	PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
	XX	PT Bower NI, Goulder KC, Green JL, Manners JM, Marcus JP;
	XX	DR WPI: 1998-377279/32.
	XX	NP Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
	XX	PS Claim 1; Page 47-49; 96pp; English.
	XX	CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian

CC	animals.
XX	
SQ	Sequence 525 AA:
	Query Match 100.0%; Score 258; DB 19; Length 525;
	Best Local Similarity 100.0%; Pred. No. 2. 7e-21;
	Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 YERDPDQQYEQCORRESEATEEREOQCQRCREREKQQQEE 47 Db 34 yerdprqeqcqqrcceseateerreqeqcqcrcereykegggqeee 80
RESULT	2
R20181	
ID	R20181 standard; Protein; 566 AA.
XX	
AC	
XX	
DE	
Macadamia integrifolia antimicrobial protein.	
XX	
KW	antimicrobial protein; infestation; control.
XX	
OS	Macadamia integrifolia.
XX	
FH	Key Location/Qualifiers
FT	Peptide 1..28 /note= "Signal peptide"
FT	Protein 29..566 /note= "mature protein"
XX	
PN	WO9827805-A1.
XX	
PD	02-JUL-1998.
XX	
PF	91WO-GB00914.
XX	
PN	WO9119801-A.
XX	
DE	Sequence encoded by 67 kd T. cacao protein cDNA.
XX	
DE	cocoa; flavour; vicilin; seed storage protein.
XX	
KW	Theobroma cacao.
XX	
OS	07-JUN-1991; 91WO-GB00914.
XX	
XX	11-JUN-1990; 90GB-0013016.
PR	
XX	26-DEC-1991.
PA	(MRSC) MARS UK LTD.
XX	
XX	WPI; 1992-024418/03.
PT	Spencer ME, Hodge R, Deakin EA, Ashton S;
PI	XX
XX	DR N-PSDB; Q20377.
XX	
PT	Recombinant cocoa proteins - are responsible for flavour in cocoa beans and produced in large quantities using yeast and bacterial expression vectors
PR	
XX	
PS	Claim 4; Fig 2; 59pp; English.
XX	
CC	The inventors claim a 67 kd and 31 kd T. cacao protein, and fragments, and encoding DNAs. The 47 kd and 31 kd proteins are derived from the 67 kd precursor. T. cacao protein cDNA was detected in a cDNA library prepared from immature cocoa beans RNA using a probe based on the AA sequence of a CbNP peptide common to the 47 kd and 31 kd polypeptides. Homology searches revealed close homologies between the 67 kd polypeptide and the vicilins, which are CC seed storage proteins.
CC	
CC	Sequence 566 AA;
SQ	
	Query Match 60.9%; Score 157; DB 19; Length 625;
	Best Local Similarity 53.3%; Pred. No. 4.6e-10;
	Matches 24; Conservative 13; Mismatches 8; Indels 0; Gaps 0;
Oy	2 ERDPDQQYEQCORRESEATEEREOQCQRCREREKQQQEE 46 ::: ::: ::: ::: ::: ::: ::: : Db 78 qrdprqqeqcqqrcceseateerreqeqcqcrcereykekkdqk 122
RESULT	4
Oy	1 YERDPDQQYEQCORRESEATEEREOQCQRCREREKQQQEE 47 Db 34 yerdprqeqcqqrcceseateerreqeqcqcrcereykegggqeee 80
RESULT	3
Db	W62829 standard; Protein; 666 AA.
XX	
AC	
XX	
DE	Macadamia integrifolia antimicrobial protein.
XX	
KW	antimicrobial protein; infestation; control.
XX	
OS	Macadamia integrifolia.
XX	
FH	Key Location/Qualifiers
FT	Peptide 1..28 /note= "Signal peptide"
FT	Protein 29..566 /note= "mature protein"
XX	
W62830	W62830 standard; Protein; 625 AA.
ID	

CC scleroses, Parkinson's disease and Leigh syndrome), cancer, cardiomypathies, ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic and mesangial disorders. The proteins may also be useful for cell growth regulation (e.g. in wound healing), T cell activation, mitosis control, and as immunosuppressants.

SQ Sequence 1233 AA;

Query Match 31.8%; Score 82; DB 20; Length 1233;
Best Local Similarity 45.7%; Pred. No. 0 18; Mismatches 10; Indels 4; Gaps 2;
Matches 21; Conservative 11; Mismatches 10; Indels 4; Gaps 2;

Qy 2 ERDPQQYEQCQRCESEATPEREQEQCERKEQORQQEE 47
||: :||: ||: ||| :||: ||: ||: :||: :||:
ID Y55931 standard; Protein: 1239 AA.
XX Y55931;

AC

XX

RESULT 15

Y55931

DT 18-FEB-2000 (first entry)

DE Human ZC1 protein.

XX

KW Antirheumatic; antithritalic; antiinflammatory; antiallergic; osteopathic; antipsoriatic; antirteriosclerotic; antiasthmatic; immunosuppressive; neuroprotective; cardiotonic; cerebroprotective; cytostatic; antidiabetic; vulneity; SMC20; protein kinase; STK2; STK3; STK4; STK5; STK6; STK7; ZC1, ZC2, ZC3, ZC4, KHS2, SUL1, SUL3, GER2, PAK1; PAK5; antagonist; antibody; gene therapy; rheumatoid arthritis; arteriosclerosis; asthma; inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis; rhinitis; autoimmunity; organ transplantation; multiple sclerosis; myocardial infarction; cardiovascular disease; stroke; renal failure; oxidative stress-related neurodegenerative disorder; Parkinson's disease; amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy; ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis; mesangial disorder; growth regulation; wound healing; T cell activation; KW immunosuppressant.

XX

OS Homo sapiens.

XX

PN W09553036 A2.

XX

PD 21-OCT-1999.

XX

PF 13-APR-1999; 99WO-US08150.

XX

PR 14-APR-1998; 98US-0081784.

PA (SUGEN INC.

XX

PT Plowman G, Martinez R, Whyte D;

XX

DR WPI; 1999-611301/52.

XX

N-PSDB; 240483.

XX

PT Novel kinase-related polypeptides used for the diagnosis and treatment of kinase-related diseases and disorders

XX

PS Claim 11; Page 269-274; 387pp; English.

XX

CC This sequence represents a novel SMC20-related protein kinase. The invention relates to nucleic acid molecule encoding a kinase polypeptide selected from STK2, STK3, STK4, STK5, STK6, STK7, ZC1, ZC2, ZC3, ZC4, KHS2, SUL1, SUL3, GEK2, PAK4 and PAK5. The proteins are used to identify agonists and antagonists, and to raise antibodies. The polynucleotides are useful in gene therapy protocols. The polynucleotides, polypeptides, antibodies, antagonists and agonists may be used to treat diseases such as immune-related disorders and diseases (e.g. rheumatoid

CC arthritis, arteriosclerosis, chronic inflammatory bowel disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplantation, chronic inflammatory pelvic disease, multiple sclerosis, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders (e.g. amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome), cancer, cardiomypathies, ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic and mesangial disorders. The proteins may also be useful for cell growth regulation (e.g. in wound healing), T cell activation, mitosis control, and as immunosuppressants.

SQ Sequence 1239 AA;

Query Match 31.8%; Score 82; DB 20; Length 1239;
Best Local Similarity 45.7%; Pred. No. 0 19; Mismatches 10; Indels 4; Gaps 2;
Matches 21; Conservative 11; Mismatches 10; Indels 4; Gaps 2;

Qy 2 ERDPQQYEQCQRCESEATPEREQEQCERKEQORQQEE 47
||: :||: ||: ||| :||: ||: ||: :||: :||:
ID 416 standard; Protein: 1239 AA.
XX 416

AC

XX

Search completed: March 1, 2001, 15:47:19
Job time: 244 sec